

- c. A hierarchical set of domain-independent meta-rules constitute a diagnostic meta-strategy. These rules examine the knowledge sources listed above and the current differential to select an hypothesis to focus on and the next datum to collect.

The key strategical idea to teach students is that collecting circumstantial evidence is preparation for making physical measurements. Its purpose is to "establish the hypothesis space," to determine the range of possibilities that might be causing the problem. Strategies for achieving this involve considering common and unusual causes, looking for evidence that will broaden the space of possibilities.

There are two orientations when establishing the hypothesis space: 1) "group and differentiate" -- upward-looking, initial problem formulation in which one tries to cluster the data under some generic process (cause); and 2) "explore and refine" -- attempting to confirm successively more specific causes.

The diagnostic meta-rules are generally applied as a pure-production system for each subtask (e.g., "find a new focus" is a subtask). Abort conditions are inherited to simulate shifting of focus (and return to higher goals) as data broaden the differential or exploration suggests that a conjecture is unlikely.

2. Research in Progress

Short-term Plans for NEOMYCIN and GUIDON

We are shifting development of GUIDON to the Dolphin computer, now on loan from Xerox and located in the Computer Science Department building at Stanford. GUIDON must be revised to be compatible with the NEOMYCIN system. These revisions take two forms: a) simplifications to the code (NEOMYCIN is designed to make it easier to index rules as they are used in the tutorial), and b) extensions to take advantage of knowledge now represented explicitly in NEOMYCIN (taxonomy of problems, world facts, diagnostic strategy).

NEOMYCIN is essentially a psychological model of diagnosis that enables us to monitor the student's problem solving and provide assistance in ways that were not possible before. For example, we will be teaching forward-directed inferences -- leaps from data to hypotheses -- that we represent in NEOMYCIN's trigger rules. With this additional knowledge of how experts think, GUIDON version 2 will have leverage for interrupting the student to test his knowledge, as well as having a better basis for understanding a student's partial solutions.

We expect that complete revisions of GUIDON so that it can take advantage of what is now in NEOMYCIN will require 6 months. This includes an entirely redesigned student model, plus the new capabilities for interruption, assistance, and evaluation of student hypotheses. In parallel, we will be refining NEOMYCIN by testing it on the 100 meningitis cases in our library. Two students will be revising GUIDON; a third student will continue development of NEOMYCIN (on the SUMEX-AIM computer facility). Dr. Clancey will direct and participate in both aspects of the project.

Formalization of Teaching Principles

One of the students who will be revising GUIDON is a doctoral candidate in the Education Department at Stanford. For his thesis research, this student will be parameterizing GUIDON's tutorial rules so they are controlled by a higher order model of teaching methods. Design of this model is complete on paper now. It will be implemented after GUIDON2 is working.

Formal Experimentation

Through our contacts with the medical school, we have arranged to test GUIDON with medical students during the period September '81 - March '82. This aspect of the project will be managed by the doctoral student in Education. Plans are to do exploratory experimentation 1) to test the usefulness of the diagnostic model for interpreting student behavior 2) determine whether theoretical differences in tutoring behavior are detectable by the students. Analysis of results should provide a basis for extending the diagnostic model.

Development of a Mechanical/Electronic Diagnostic Program

We have begun collaboration with researchers from IBM to develop a system similar to NEOMYCIN in the domain of computer failure diagnosis. The purpose of this project will be to determine to what extent the domain-independent strategies we formalized from experience in the medical domain are applicable to electronic troubleshooting. In the past year, Prof. Buchanan supervised development of an EMYCIN consultation program, named DART, for diagnosing teleprocessing problems. Through this experience, IBM personnel learned about our techniques, and we were introduced to the hardware and software problems they need to solve. We will be drawing upon this experience in the next year.

ONCOCIN

The oncology chemotherapy consultation system, named ONCOCIN, has achieved many of its goals since work on the project began in July 1979. We are developing an interactive system to be used by oncology faculty and fellows in the Debbie Probst Oncology Day Care Center at Stanford University Medical Center. Our overall goals are:

- (1) to demonstrate that a rule-based consultation system with explanation capabilities can be usefully applied and gain acceptance in a busy clinical environment;
- (2) to improve the tools currently available, and to develop new tools, for building knowledge-based expert systems for medical consultation; and
- (3) to establish both an effective relationship with a specific group of physicians, and a scientific foundation, that will together facilitate future research and implementation of computer-based tools for clinical decision making.

Specific Objectives:

The ONCOCIN research goals are directed both towards the basic science of artificial intelligence and towards the development of clinically useful oncology consultation tools.

Artificial Intelligence Objectives

We have undertaken AI research with the following aims:

- (1) to implement and evaluate recently developed techniques designed to make computer technology more natural and acceptable to physicians;
- (2) to extend the methods of rule-based consultation systems to interact with a large database of clinical information; and
- (3) to continue basic research into the following problem areas: mechanisms for handling time relationships, techniques for quantifying uncertainty and interfacing such measures with a production rule methodology, approaches to acquiring knowledge interactively from clinical experts, assessment of knowledge base completeness and consistency.

Oncology Clinic Objectives

We have begun to develop and implement a protocol management system, for use in the oncology day care center with the following capabilities:

- (1) to assist with identification of current protocols that may apply to a given patient;
- (2) to assist with determining a patient's eligibility for a given protocol;
- (3) to provide detailed information on protocols in response to questions from clinic personnel;
- (4) to assist with chemotherapy dose selection and attenuation for a given patient;

- (5) to provide reminders, at appropriate intervals, of follow-up tests and films required by the protocol in which a given patient is enrolled;
- (6) to reason about managing current patients in light of stored data from previous visits of (a) the individual patients, or (b) the aggregate of all "similar" patients.

Overview of Goals for 1980:

We have described a five-year plan for accomplishing the above goals. As discussed at this time last year, we spent our first year developing a prototype ONCOCIN consultation system, drawing from programs and capabilities developed for the EMYCIN system-building project. During that year, we also undertook a detailed analysis of the day-to-day activities of the Stanford oncology clinic in order to determine how to introduce ONCOCIN with minimal disruption of an operation which is already running smoothly. We also spent much of our time in the first year giving careful consideration to the most appropriate mode of interaction with physicians in order to optimize the chances for ONCOCIN to become a useful and accepted tool in this specialized clinical environment.

During our second year of the project, we have accomplished all the goals we identified for 1980:

- (1) We have completed a special interface program that responds to commands from the customized keypad described last year;
- (2) We encoded the rules for one more chemotherapy protocol (oat cell carcinoma of the lung) and updated the Hodgkin's Disease protocols when new versions were released late in 1980; these exercises demonstrated the generality and flexibility of the representation scheme we have devised;
- (3) We developed the software protocols for achieving communication between the interface program and the reasoning program;
- (4) We have coordinated the printing routines needed to produce hardcopy flowsheets, patient summaries, and encounter sheets;
- (5) Lines have been installed between the SUMEX machine room and the oncology clinic, and the new terminal and a hard copy device have been installed in the Oncology Day Care Center for final testing and debugging; and
- (6) We have just begun to offer the ONCOCIN system for use by oncology faculty and fellows in the morning chemotherapy clinics in which most of the lymphoma patients receive their treatment.

We had two additional goals, not explicitly stated in last year's report. One was to design formal evaluation studies that would allow us to assess the impact of ONCOCIN and its acceptance by the physicians for whom

it is designed. Second, we wanted to experiment with computational techniques for verifying the completeness and consistency of a developing knowledge base.

PROGRESS - 1980/81:

Further Development and Testing of the Reasoning Program

The early prototype of the Reasoning system was described in last year's report in some detail. A more recent summary has been submitted for presentation at the 7th International Joint Conference on Artificial Intelligence. The Reasoner is coded in Interlisp, and is running on the SUMEX computers (both the PDP-10, and the 20/20 on which we have been running when the system is used in the oncology clinic).

The Reasoner has been extensively debugged this year. Several hundred sample patient cases have been run, and the results have been reviewed in detail by the collaborating oncologists. When problems have been uncovered by this process, changes in the Reasoner program (or in the encoding of the lymphoma protocol knowledge) have been undertaken.

Verification of the Adequacy of the Knowledge Representation Scheme

In an effort to verify that the representation scheme we are using will be adequate for arbitrary protocol knowledge that may be encountered in the future, we decided to encode and briefly test the knowledge of a non-lymphoma protocol. We chose the complicated protocol for oat cell (small cell) carcinoma of the lung because it involves a large number of possible therapies and complex interweaving of chemotherapy and radiotherapy. After approximately one month's effort by an experienced programmer, the oat cell protocol had been encoded and ran successfully on a few test cases. In addition, the lymphoma protocols themselves were changed in late 1980, and we spent a few weeks in early 1981 entering the changes implicit in these new versions. In all cases the ONCOCIN representation scheme was adequate to accommodate the protocol knowledge with only minor modifications, if any, and for this reason we are confident that our system will be able to adapt to any other protocols that need to be encoded in the coming years.

Physician/Computer Interaction

The actual mechanics of computer terminal interaction is as important to a clinical system's acceptance as the quality of the program's advice. If a system is slow or cumbersome, physicians will tend to reject it. With this in mind, we have sought to develop an optimal interactive mechanism that will not unreasonably tax the budget of the project.

In last year's report we indicated that this interactive system was to be written in PASCAL. After some initial experiments, however, we decided to use SAIL instead. The system is referred to as the "Interviewer", and it has now been fully implemented and debugged.

As we emphasized when outlining our research goals, we have wanted ONCOCIN to maintain the explanation and justification capabilities that we have argued are crucial to the acceptance of clinical consultation systems. The Interviewer uses a specialized split-screen display that enables the physician to enter patient data entries in one region while pertinent explanations are displayed in another.

Development of Mechanisms for Interprocess Communication

Because the Reasoner (the Interlisp reasoning program) and the Interviewer (the SAIL program with which the physician interacts) must run in parallel in two different processes on the same machine, we needed to devise mechanisms for allowing these two programs to communicate with one another. This has been a major systems programming task, but we are pleased with the effectiveness of the generalized interprocess communication mechanism that we devised.

Designing an Evaluation of the ONCOCIN System

Because we wish to evaluate formally the impact of ONCOCIN and its effectiveness in the oncology clinic, we have devised a set of three experiments, two of which are already underway. The study designs are outlined in detail in an evaluation document that we have prepared.

Verifying the Completeness and Consistency of the Knowledge Base

An important question for AI researchers involved with the development of expert systems is how to ascertain that a knowledge base for a consultation program is complete and consistent. Dr. Motoi Suwa, a visitor to Stanford from Japan, became fascinated with this question and collaborated with us on a formal analysis of the developing ONCOCIN knowledge base. His paper describing that work was submitted for presentation at the 7th International Joint Conference on Artificial Intelligence.

D. Publications Since January 1980

- Aikins, Janice S. Prototypes and Production Rules: A Knowledge Representation for Computer Consultations. PhD Dissertation, Stanford University. Memo HPP-80-17, August, 1980.
- Bennett, S.W., and Scott, A.C. Computer-assisted customized antimicrobial dosages. Amer. J. Hosp. Pharm. 37:523-9 (1980).
- Buchanan, B.G. Research on Expert Systems. Memo HPP-81-1, January, 1981. To appear in D. Michie (ed.) Machine Intelligence 10.
- Campbell, A.B., Chang, P., Cho, J., Hickam, D., Shortliffe, E.H., and Teach, R. Preliminary proposal for the evaluation of the ONCOCIN System. Internal memo, the ONCOCIN Project, April 1981.

- Clancey, W.J. The Epistemology of a Rule-based Expert System: the effect of proceduralization of knowledge on explanation. To appear in the Journal of Artificial Intelligence.
- Clancey, W.J. and Letsinger R. NEOMYCIN: Reconfiguring a rule-based expert system for application to teaching. Submitted to IJCAI7.
- Clancey, W.J. Methodology for building an intelligent tutoring system. To appear in a book on Cognitive Science Methodology, edited by Kintsch, Miller, and Polson.
- Fagan, L.M., Shortliffe, E.H., and Buchanan, B.G. Computer-based medical decision making: from MYCIN to VM. Automedica 3:97-106 (1980).
- Gerring, Phil. System documentation: interprocess communication system (TopDog and Interactor). Internal memo, the ONCOCIN Project, November 1980.
- Shortliffe, E.H., Scott, A.C., Bischoff, M.B., Campbell, A.B., van Melle, W., and Jacobs, C.D. ONCOCIN: An expert system for oncology protocol management. Submitted to Proceedings of the 7th IJCAI, Vancouver, B.C., August 1981.
- Shortliffe, E.H. Consultation systems for physicians. Proceedings of the CSCSI/SCEIO Conference, 14-16 May 1980, University of Victoria, British Columbia, pp. 1-11.
- Shortliffe, E.H. Medical Cybernetics: The Challenges of Clinical Computing. To appear in Cybernetics, Technology, and Growth, S. Basheer Ahmed, editor; Lexington Books, 1981.
- Shortliffe, E.H. Medical Computing: Another Basic Science? Proceedings of the 4th Symposium on Computer Applications in Medical Care, Washington, D.C., November 1980.
- Shortliffe, E.H. The computer as clinical consultant (editorial). Arch. Int. Med., March 1980.
- Suwa, M., Scott, A.C., and Shortliffe, E.H. An approach to verifying completeness and consistency in a rule-based expert system. Submitted to Proceedings of 7th IJCAI, Vancouver, B.C., August 1981.
- Teach, R.L. and Shortliffe, E.H. An analysis of physician attitudes regarding computer-based clinical consultation systems. Submitted for publication in Comput. Biomed. Res., February 1981.
- van Melle, W., Shortliffe, E.H., Buchanan, B.G. EMYCIN: A domain-independent system that aids in constructing knowledge-based consultation programs. To appear in Pergamon-Infotech State of the Art Report on Machine Intelligence, 1981.
- van Melle, W. A domain-independent system that aids in constructing knowledge-based consultation programs. PhD thesis, Computer Science Department, Stanford University, June, 1980.

E. Funding Support

Grant Title: "Knowledge-Based Consultation Systems"

Principal Investigator: Bruce G. Buchanan

Agency: National Science Foundation

ID Number: MCS-7903753

Term: July 1979 to March 1981

Total award: \$146,152

Current award (1980): \$72,493

[No continuation proposal was submitted to the NSF since the current version of the system successfully completes our proposed work. We intend to use EMYCIN as a vehicle for experimental research under other funding, including SUMEX core research, but we are not proposing further research or development on EMYCIN itself.]

Contract Title: "Exploration of Tutoring and Problem-Solving Strategies"

Principal Investigator: Bruce G. Buchanan

Agency: Office of Naval Research and
Advanced Research Projects Agency (joint)

ID number: N00014-79-C-0302

Term: March 1979 to March 1982

Total award: \$396,325

Grant Title: "Explanatory Patterns In Clinical Medicine"

Principal Investigator: Edward H. Shortliffe

Agency: Kaiser Family Foundation

Term: July 1979 to December 1980

Total award: \$20,000

SKIP 2

Grant Title: "Research Program: Biomedical Knowledge Representation"

Principal Investigator: Edward A. Feigenbaum

Co-Principal Investigator (ONCOCIN Project): Edward H. Shortliffe

Agency: National Library of Medicine

ID Number: LM-03395

Term: July 1979 to June 1984

Total award: \$497,420

Current award (1980-1981): \$99,400

Administered through Medicine: ONCOCIN suballocation (\$47,845)

Grant Title: "Symbolic Computation Methods For Clinical Reasoning"

Principal Investigator: Edward H. Shortliffe

Agency: National Library of Medicine

ID Number: LM-00048

Term: July 1979 to June 1984

Total award: \$196,425

Current award (1980-1981): \$39,107

II. INTERACTION WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

A great deal of interest in both MYCIN and EMYCIN has been shown by the medical and academic communities. For two years in succession we were invited by the American College of Physicians to demonstrate MYCIN at the organization's annual meeting (San Francisco, March 1979, and New Orleans, April 1980). The physicians have uniformly been enthusiastic about the program's potential and what it reveals about one current approach to computer-based medical decision making. In both cases, the demonstrations were performed on-line using network access to the SUMEX computer.

We have demonstrated our programs to both physicians and computer scientists on numerous additional occasions. At the AIM tutorial in August 1980, both MYCIN and GUIDON were presented to introduce physicians to the field of AI in medicine. GUIDON was also demonstrated on the Dolphin machine at the Xerox-PARC open house during the AAAI in August. In addition, both EMYCIN and GUIDON were featured demonstrations at the annual AIM Workshop, held the same week at Stanford. The TYPER program, developed by SUMEX staff in collaboration with Dr. Larry Fagan of Stanford, was used to good effect at this workshop as well as for informal demonstrations throughout the year.

Several project members contributed to the Expert Systems Workshop, sponsored by RAND and ARPA and held in San Diego in August, where EMYCIN was one of the "system building tools" that was studied in detail. The Workshop has led to the preparation of a book, "Building Expert Systems," and many of our research group have written portions of that volume (Buchanan, Clancey, Scott, Aikins, Shortliffe, van Melle).

NEOMYCIN was presented to the contractors of the ONR "Instructional systems and advanced training" division, held in Pittsburgh, in January 1981. Presentations of this kind carry SUMEX-AIM results out to cognitive psychologists from around the country. Dr. Clancey also presented a talk on GUIDON research at the annual conference of the Association for Development of Computer Instructional Systems in Atlanta, in March 1981, and at the annual conference of AERA in Los Angeles, in April 1981.

Several medical school and computer science teachers have also asked to use MYCIN in their computer science or medical computing courses, and we continue to make the programs available frequently to researchers around the world who access SUMEX using the GUEST account.

EMYCIN has generated considerable interest in the academic and business communities. We have been in frequent contact with Bud Frawley and Alain Bonnet, of Schlumberger, Chuck Brodnax and Milt Waxman of the Hughes Aircraft Corporation, and Harry Reinstein and Cliff Hollander from IBM Scientific Research Center. EMYCIN, on SUMEX, has been used at the University of Illinois and Michigan State University to explore the construction of expert systems.

B. Sharing and Interaction with Other SUMEX-AIM Projects

We have continued collaboration with the RX, VM, and PUFF projects. Our development of a domain-independent system is facilitated by having a number of very different working systems on which to test our additions and modifications to EMYCIN. All the projects have provided us with useful comments and suggestions.

The community created on the SUMEX resource has other benefits that go beyond actual shared computing. Because we are able to experiment with other developing systems, such as INTERNIST, and because we frequently interact with other workers (at the AIM Workshop or at other meetings around the country), many of us have found the scientific exchange and stimulation to be heightened. Several of us have visited workers at other sites, sometimes for extended periods, in order to pursue further issues which have arisen through SUMEX- or Workshop-based interactions. In this regard, the ability to exchange messages with other workers, both on SUMEX and at other sites, has been crucial to rapid and efficient exchange of ideas. For example, most of the invitations and planning for the 6th AIM Workshop, held at Stanford in August 1980 and described in detail elsewhere in this report, were accomplished via SUMEX or ARPANET mail. Certainly it is unusual for a small community of researchers with similar scholarly interests to have at their disposal such powerful and efficient communication mechanisms, even among those on opposite coasts of the country.

C. Critique of Resource Management

The SUMEX facility has maintained the high standards that we have praised in the past. The staff members are always helpful and friendly, and work as hard to please the SUMEX community as to please themselves. As a result, the computer is as accessible and easy to use as they can make it. More importantly, it is a reliable and convenient research tool. We extend special thanks to Tom Rindfleisch for maintaining high professional standards for all aspects of the facility.

Due to the introduction of our ONCOCIN work with its special hardware and communication needs, we continue to be aware that we are taxing the limited resources of SUMEX with regards to technical hardware support. It has been next to impossible for one technical specialist (Nick Veizades) to balance the numerous diverse demands on his time. This is not a problem with management of the Resource but a reflection of the need for additional technical personnel associated with SUMEX. We perceive this to be a particularly important requirement in the future as the Resource undertakes an expanded role in the implementation and testing of new hardware.

Special mention should be made of the remarkable role played by Tom Rindfleisch and his staff in helping to organize remote demonstrations of SUMEX-AIM programs. In October, 1980, when the NIH Council on Research Resources met in Atlanta, demonstrations of MYCIN and INTERNIST on the DEC 2020 at Stanford were so carefully arranged as to make them seem commonplace. We salute Tom and the staff for their uncomplaining assistance, and are grateful for the efforts they have made to provide a mechanism for facilitating future demonstrations at remote locations.

Finally, we continue to feel the need for more computing power. Much of our research and development continues to take place in the hours from 7 p.m. to 10 a.m., but it is unreasonable to expect all our programming staff to adjust their own schedules around a computer. The existence of the 20/20 has been helpful in permitting demonstrations with good response time, and it has allowed us to introduce ONCOCIN in a real clinical environment, but ongoing R&D on the main machine remains difficult much of the time. Even the evening hours are now seeing higher load averages than was once the case. We anticipate considerable improvement in this regard as the recently approved additional computing hardware becomes available. In the meantime, much of the work on EMYCIN has been moved to the SCORE computer in the Computer Science Department. Response time aside, we have shifted our development of GUIDON to the Xerox Dolphin in order to take advantage of the larger address space. This also frees up disk space so that we can comfortably develop NEOMYCIN on SUMEX.

We also strongly support the creation of the new position assumed by Anne Fadenrecht; her excellent early efforts should be especially helpful in taking some of the load off of Carole Miller and Tom Rindfleisch.

III. RESEARCH PLANS

A. Project Goals and Plans

EMYCIN

Now that the design and capabilities of EMYCIN are essentially fixed, we are planning to develop new applications and to use the system as an experimental tool. The applications to electronic fault diagnosis and geology will continue and we expect to find additional medical applications as well. Many of these we expect will be undertaken by other research groups. Because we view artificial intelligence as an experimental science [Buchanan, 1981], we wish to collect data on the nature of problems EMYCIN can help solve and the limitations of the problem solving method embodied in EMYCIN.

Our research on knowledge acquisition depends on the existence of a working EMYCIN system. In the ROGET program, currently under development, hierarchical knowledge about consultation systems and their knowledge bases is used to help an expert define a new knowledge base to be used by EMYCIN. For example, the meningitis and pulmonary function knowledge bases both contain rules associating diagnoses with laboratory tests and with clinical findings. ROGET will be able to use this fact to help an expert divide a new rule set into rules using test results and measurements as evidence and another rule set using more subjective evidence.

GUIDON

We have now established a good framework for organizing knowledge in an expert system to be used for tutoring. We characterize knowledge by its use for: structuring knowledge sources, supporting (justifying) knowledge sources, or controlling their invocation. In the most general terms, our plans are to do research in acquiring, representing and presenting structural, support, and strategic knowledge.

We used this framework to design NEOMYCIN. Experiments during the coming year will provide a basis for developing our model of diagnosis. In particular, we propose:

- a) to extend NEOMYCIN's model of diagnostic strategy to include common, non-expert approaches. Besides improving the program's ability to model the student, this enumeration of the space of strategies will allow us to follow a plan of research similar to Brown's and Burton's, but in the domain of diagnostic strategy as opposed to subtraction procedures. Eventually, we want to develop a principled psychological model that will relate strategies to knowledge and processing abilities.
- b) further studies of expert reasoning in domains that require "forming a picture" of a malfunctioning process. Experience with NEOMYCIN showed that expert diagnosticians attempt to order the data they collect causally, on a time line. Interpretation of observations can be partially understood as an attempt to match this description of onset, course, severity (intensity, frequency), and causal relations of findings onto known malfunctions that are recalled (indexed) by these process variables. This work will build upon recent advances in understanding causality (e.g., deKleer and Brown).
- c) exploitation of new technology for experimentation with teaching methods. How can we take advantage of the Dolphin's graphic capabilities in a GUIDON tutorial? Besides graphically presenting rule relationships, we might show the student the same kind of diagrams that we use when describing our knowledge bases to our AI colleagues (hierarchies, diagrams relating compiled associations to underlying causal chains). Other than presentation strategies, we would like to experiment with different interfaces, perhaps to break away from a continuous dialogue to use the screen more as a work space for annotating and examining the knowledge base, and organizing data and hypotheses in a diagnostic problem.

- d) incorporate GUIDON as an integral part of the curriculum in in medical diagnosis at Stanford. We propose to make GUIDON available at the Fleischmann Learning Center at Stanford Medical School, just as the traditional programs built at Massachusetts General and Ohio State were made available. In addition, we will work with one or more teaching fellows at the medical school to include GUIDON as part of the "clinical diagnosis" course which is taught regularly at Stanford. This will continue our commitment to empirical research to develop our model of diagnosis and the teaching procedures.

ONCOCIN

During the coming year, there are four principal areas in which we expect to expend our efforts on the ONCOCIN System:

- (1) The system will be implemented for ongoing use in the Stanford Oncology Clinic, with an experimental evaluation period to begin July 1, 1981.
- (2) The system will be formally evaluated with regard to its impact on (a) the attitudes of the oncologists, (b) the accuracy and completeness of data collection, and (c) the adequacy of the management decisions made in the clinic.
- (3) We will begin to encode additional protocols as the lymphoma system comes into regular use and physicians begin to demand the inclusion of a greater percentage of the protocols used in the management of cancer patients at Stanford.
- (4) We will begin to devote a greater percentage of our time to experiments in encoding complex judgmental reasoning of the sort that is usually performed by expert oncologists and is not formally specified in the protocol documents themselves.

Throughout the year we shall continue to relate the requirements of the system we are developing to the underlying artificial intelligence methodologies. We are convinced that the basic science frontiers of AI are best explored in the context of systems for real world use; thus ONCOCIN serves as a vehicle for developing an improved understanding of the issues that underlie all forms of knowledge engineering.

B. Requirements for Continued SUMEX Use

All the work we are doing (EMYCIN, GUIDON, ONCOCIN, plus continued use of the original MYCIN program) is totally dependent on continued use of the SUMEX resource. The programs all make assumptions regarding the computing environment in which they operate, and the ONCOCIN design in particular depends upon proximity to the DEC 2020 which enables us to use a 9600 baud interface.

In addition, we have long appreciated the benefits of GUEST and network access to the programs we are developing. SUMEX greatly enhances our ability to obtain feedback from interested physicians and computer scientists around the country. Network access has also permitted high quality formal demonstrations of our work both from around the United States and from sites abroad (e.g., Japan, Sweden, Switzerland).

We plan to continue development of NEOMYCIN on SUMEX during the next year, whereas the GUIDON/Dolphin effort will continue on the crowded Computer Science Department Dolphin only until the SUMEX individual workstations become available. Using the main SUMEX machine, we intend to make NEOMYCIN fully usable as a consultation program so that it can be compared with MYCIN. In particular, we will be comparing cases run through both MYCIN and NEOMYCIN to see whether simplification and clarification of the rules for purposes of teaching will in turn change the program's accuracy.

C. Requirements for Additional Computing Resources

The acquisition of the DEC 2020 by SUMEX has been crucial to the growth of our research work, both to insure high quality demonstrations and to enable us to develop a system such as ONCOCIN for real-world use in a clinical setting. As we continue to develop systems that are potentially useful as stand-alone packages (e.g., an exportable EMYCIN), the additional small computers that are planned will be particularly valuable resources. It is not yet clear which machines are optimal for the LISP-based applications we are developing, and an opportunity to test our systems on several small-to-medium machines will be invaluable and in keeping with our desire to move some of the AIM products into a community of service users.

As we have mentioned, the response time on the main machine continues to be a major problem, both during the daytime hours and frequently in the evenings as well. The proposed SUMEX acquisitions that will provide additional cycles and permit off-loading of some users from the PDP-10 will significantly benefit the SUMEX research community.

In addition, we believe that our GUIDON experience using the Dolphin personal computer is a significant part of our research. First, the Dolphin's large address space will permit development of the large knowledge base that an intelligent tutoring system requires; we have overgrown the facilities available at SUMEX. Second, the Dolphin's graphics will enable us to develop new methods for presenting material from the knowledge base. Third, the Dolphin will provide a reliable, constant "load-average" machine, for running experiments with students. Finally, the development of GUIDON on the Dolphin demonstrates the feasibility of running intelligent tutoring systems on small, affordable machines in schools and remote sites.

We seem to have an insatiable appetite for disk storage space, even though ONCOCIN received an additional substantial allocation since our report last year. ONCOCIN, in particular, has become an extremely large system, and the data files for a clinic full of patients will require substantial additional space. We hope that the planned SUMEX file-server

will allow the allocation of several thousand more pages. It should also help alleviate the need to keep copies of all patient files on both the 10 and the 20/20.

D. Recommendations for Future Community and Resource Development

In last year's report we made two recommendations for new SUMEX developments: (1) the acquisition of several small machines, linked to the main processor through the Ethernet and able to run INTERLISP, and (2) the formal establishment of a mechanism for providing hardware and communications equipment for SUMEX demonstrations at a distance. Both of these have been acted upon by SUMEX, and we are delighted by this kind of responsiveness.

The AIM community is small and close-knit, but there remain communication problems within it. The AIM Workshops are excellent means of transferring information annually, but between Workshops all of us are remiss in not communicating new technical reports and articles. It would be very desirable to maintain a list of current publications from all the AIM research groups, for distribution by ARPANET or U.S. Mail to all others. No group will add to the list, however, unless the benefit of the information gained from such a list exceeds the cost of adding to it. SUMEX may be able to function as a catalyst to this kind of community communication.

II.A.1.7 Protein Structure Project

Protein Structure Modeling Project

Prof. E. Feigenbaum and Mr. Allan J. Terry
Department of Computer Science
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Technical Goals

The goals of the protein structure modeling project are to 1) identify critical tasks in protein structure elucidation which may benefit by the application of AI problem-solving techniques, and 2) design and implement programs to perform those tasks. We have identified two principal areas which are of practical and theoretical interest to both protein crystallographers and computer scientists working in AI. The first is the problem of interpreting a three-dimensional electron density map. The second is the problem of determining a plausible structure in the absence of phase information normally inferred from experimental isomorphous replacement data. Current emphasis is on the implementation of a program for interpreting electron density maps (EDM's).

B. Medical Relevance and Collaboration

The biomedical relevance of protein crystallography has been well stated in an excellent textbook on the subject (Blundell & Johnson, Protein Crystallography, Academic Press, 1976):

"Protein Crystallography is the application of the techniques of X-ray diffraction ... to crystals of one of the most important classes of biological molecules, the proteins. ... It is known that the diverse biological functions of these complex molecules are determined by and are dependent upon their three-dimensional structure and upon the ability of these structures to respond to other molecules by changes in shape. At the present time X-ray analysis of protein crystals forms the only method by which detailed structural information (in terms of the spatial coordinates of the atoms) may be obtained. The results of these analyses have provided firm structural evidence which, together with biochemical and chemical studies, immediately suggests proposals concerning the molecular basis of biological activity."

The project involves a collaboration between computer scientists at Stanford University and crystallographers at Oak Ridge National Laboratories (Dr. Carroll Johnson), the University of California at San

Francisco (Dr. Robert Langridge), and the University of California at San Diego (under the direction of Prof. Joseph Kraut). Our principal collaborator at UCSD is Dr. Stephan Freer. We also collaborate with Dr. Eric Grosse at Bell Laboratories, whose field is numerical analysis.

C. Progress Summary

We have completed a major cycle of design review and program reorganization, resulting in the system described in publications four and nine below. The system now has a completely hierarchical, rule-based control structure proceeding from strategy rules, to a set of task rules, ending with individual knowledge sources. This new design seems powerful and flexible enough to provide the basis of a useful EDM interpretation system for protein structure determination.

After building the control structure we wanted, we have worked on building up the knowledge base. Large chunks of knowledge are called "tasks"; we have implemented five out of a projected set of nine. To date, we have implemented the Initialization task, two tracing tasks, a task to split "group toeholds", and a version of a task that finds "second generation" toeholds. Further details of these tasks and their content can be found in publication number four.

We have also continued our efforts to improve the power of our data representations. Towards this end we have implemented a new preprocessor based on Dr. Grosse's thesis research. This program is an improved method for finding the critical points of a function. In our case, the peaks of the electron density map are useful guides to atom locations and the full set of critical points are used in the ridge-line analysis discussed in publication one.

Finally, we are compiling documentation on the system and the knowledge it embodies. These documents should be sufficiently complete so that we, or other groups, will have little difficulty picking up where we leave off. We also feel that explicit documentation of our model-building heuristics will be useful to the crystallographic community as it provides a new viewpoint, complementary to traditional crystallographic methods.

The work currently in progress can be characterized as additions to the knowledge base and work on new data representations. The five tasks currently implemented form the core of the system and suffice to solve about a half of a small protein. The remaining tasks will embody knowledge about finding new toeholds (to restart the trace when it is blocked) and about tracing in areas of the data too complex to interpret with present heuristics. One of the main areas of work along these lines is the incorporation of some notion of stereochemistry and the constraints on three dimensional structure it provides. This will be useful in the matching of features and in the prediction of secondary structure. The last item of work in progress is an attempt to design a data representation that captures volume information. Current representations such as the skeleton preserve topology but do not preserve shape. With the inclusion of volume information, we should be able to capture much of the expert's knowledge of shape and form that presently goes unused.

D. List of Publications

- (1) Carroll Johnson and Eric Grosse, "Interpolation Polynomials, Minimal Spanning Trees, and Ridge-Line Analysis in Density Map Interpretation", American Crystallographic Association Program and Abstracts, 4:2, Evanston, Ill. Aug. 1976
- (2) Robert S. Engelman and H. Penny Nii, "A Knowledge-Based System for the Interpretation of Protein X-Ray Crystallographic Data," Heuristic Programming Project Memo HPP-77-2, January, 1977. (Alternate identification: STAN-CS-77-589)
- (3) E.A. Feigenbaum, R.S. Engelman, C.K. Johnson, "A Correlation Between Crystallographic Computing and Artificial Intelligence," in Acta Crystallographica, A33:13, (1977). (Alternate identification: HPP-77-15)
- (4) Robert Engelman and Allan Terry, "Structure and Function of the CRYNALIS System", Proc. 6IJCAI, 1979. pp250-256 (Alternative identification: HPP-79-16)
- (5) R.S. Engelman, A. Terry, S.T. Freer, and C.K. Johnson, "A Knowledge-Based System for Interpreting Protein Electron Density Maps", Abstracts of Amer. Crystallographic Ass. 7,1 (1979) p38
- (6) E.H. Grosse, "Approximation and Optimization of Electron Density Maps", Stanford University Ph.D. Thesis, Dec. 1980 (Alternative identification: STAN-CS-80-835)
- (7) R. Engelman and A. Terry, Article VII.C3 (Crysalis) in Barr, A., and Feigenbaum, E. A. (eds.), The Handbook of Artificial Intelligence, Vol. II, Stanford Ca., HeurisTech Press, Los Altos, Ca.: Kaufman, 1981
- (8) A. Terry and R. Engelman, "A Knowledge-Based Approach to the Interpretation of Protein Electron Density Maps", to appear in a forthcoming book on expert systems by Pergamon Infotech International, Maidenhead, England
- (9) A. Terry, "Hierarchical Control of Production Systems", paper submitted to 7IJCAI

E. Funding status

Grant title: The Automation of Scientific Inference: Heuristic
Computing Applied to Protein Crystallography

Principal Investigator: Prof. Edward A. Feigenbaum

Funding Agency: National Science Foundation

Grant identification number: MCS 79-33666

Term of award: December 1, 1979 through November 31, 1981

Amount of award: \$35,318 (direct costs only)

II. INTERACTION WITH THE SUMEX-AIM RESOURCE

A. Collaborations

The protein structure modeling project has been a collaborative effort since its inception, involving co-workers at Stanford and UCSD (and, more recently, at Oak Ridge, UCSF, and Bell Laboratories). The SUMEX facility has provided a focus for the communication of knowledge, programs and data. Without the special facilities provided by SUMEX the research would be seriously impeded. Computer networking has been especially effective in facilitating the transfer of information. For example, the more traditional computational analyses of the UCSD crystallographic data are made at the CDC 7600 facility at Berkeley. As the processed data, specifically the EDM's and their Fourier transforms, become available, they are transferred to SUMEX via the FTP facility of the ARPA net, with a minimum of fuss. (Unfortunately, other methods of data transfer are often necessary as well -- see below.) Programs developed at SUMEX, or transferred to SUMEX from other laboratories, are shared directly among the collaborators. Indeed, with some of the programs which have originated at UCSD and elsewhere, our off-campus collaborators frequently find it easier to use the SUMEX versions because of the interactive computing environment and ease of access. Advice, progress reports, new ideas, general information, etc. are communicated via the message and/or bulletin board facilities.

B. Interaction with Other SUMEX-AIM Projects

Our interactions with other SUMEX-AIM projects have been mostly in the form of personal contacts. We have strong ties to the MYCIN, AGE and MOLGEN projects and keep abreast of research in those areas on a regular basis through informal discussions. The SUMEX-AIM workshops provide an excellent opportunity to survey all the projects in the community. Common research themes, e.g. knowledge-based systems, as well as alternate problem-solving methodologies were particularly valuable to share.

C. Critique of Resource Services

The SUMEX facility provides a wide spectrum of computing services which are genuinely useful to our project -- message handling, file management, Interlisp, Fortran and text editors come immediately to mind. Moreover, the staff, particularly the operators, are to be commended for their willingness to help solve special problems (e.g., reading tapes) or providing extra service (e.g. immediate retrieval of an archived file). We would also like to commend the staff for its extensive help in setting up a link between SUMEX and Dr. Langridge's group at UCSF. Such cooperative behavior is rare in computer centers.

There are several facilities we wish to single out as particularly useful in furthering our research goals. Since the members of the project are physically distant, the MSG program is very useful. Similarly, the file system, the ARCHIVE facility, and the general ease of getting backup files from the operator greatly aid our efforts at coordinating the efforts of collaborators using many large data sets and programs. The crystallographers in the project find SUMEX to be a friendly environment which allows them to do their work with a minimum of dealing with operating system details.

It has become increasingly evident, however, that as CRYNALIS expands, the facility cannot provide enough machine cycles during prime time to support the implementation and debugging of new features. For example, our segment-labeling preprocessor requires about an hour of machine time per 100 residues of protein (this is typically five to eight hours of terminal time during working hours) even when the Lisp code is compiled.

III. USE OF SUMEX DURING THE REMAINING GRANT PERIOD (8/79 - 7/81)

A. Long-Range Goals

Our short term goals are to build up the knowledge base to the point where it can solve a small, known protein from "live" data. This will probably entail the implementation of at least seven tasks. By this point we should also have a package of data-reduction programs suitable for export to interested crystallographers.

Our long range goals are the exploitation of the rule-based control structure for investigating alternative problem-solving strategies, the investigation of modes of explanation of the program's reasoning steps, and the expansion and generalization of the system to cover a wider range of input data.

B. Justification for Continued Use of SUMEX

We feel that SUMEX is the ideal vehicle for further research on CRYNALIS. While some of our work is numerical in nature and uses such facilities as FORTRAN, our main interest is in artificial intelligence. Besides being an expert system of use to the crystallographic community,

CRYNALIS is an exploration of the general signal processing problem. We are vitally concerned with issues such as proper architecture for using a wide variety of heuristics effectively and hypothesis formation when both data and model are poor. The utility of our work to the AI community is partially demonstrated by the development of the AGE project, an extension of Ms. Nii's early work on CRYNALIS.

This project progresses by the collaboration of several physically-separated groups. SUMEX provides a unique resource, an electronic community of researchers in our field, through the many systems such as net mail, country-wide access, and community workshops. We feel that CRYNALIS would not be possible outside of such a community.

C. Needs and Plans for Other Computing Resources

Our major need for other computing resources is for graphical display of our data and results. This need will be met by use of Dr. Langridge's Evans and Sutherland Picture System at UCSF and Dr. Johnson's raster-based graphics system at ORNL. The major impediment is SUMEX's current inability to support data transfer to other machines at more than 1200 baud. We are attempting to link SUMEX to UCSF by using FTP over the ARPAnet to the LBL machine and then use an existing link from LBL to UCSF.

We will make minor use of the Stanford Computer Science Department's SCORE machine, mostly to run the SCRIBE text formatting program until such time as it is available on SUMEX.

D. Recommendations for Future Community and Resource Development

There are two recommendations we wish to make, the first and most important is to expand the computing power available to SUMEX users. CRYNALIS is an inherently-large problem. Proteins contain hundreds, to thousands of atoms which means large hypothesis structures, large quantities of data, and a compute-bound inference program. As the system grows to maturity, we expect increasingly serious problems with address space limitations and with machine cycle availability.

The second recommendation is that SUMEX develop some relatively inexpensive file transfer facility for machines not on the ARPAnet. Software for this already exists in the form of the TTYFTP program (or possible future programs like it, but in a more portable language), the development needed is in hardware and in the TENEX operating system so that transfer rates greater than 1200 baud can be achieved. We are motivated to recommend this not only by our own need for such a facility, but also by the belief that it would aid other collaborations involving SUMEX and outside computers (the SECS project for example), and aid in the dissemination of useful programs from the research setting of SUMEX to user laboratories.

II.A.1.8 RX ProjectThe RX Project: Deriving Medical Knowledge from
Time-Oriented Clinical Databases

Robert L. Blum, M.D.
Department of Computer Science
Stanford University

Gio C. M. Wiederhold, Ph.D.
Departments of Computer Science and Electrical Engineering
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Technical Goals

Introduction:

Medical and Computer Science Goals

The objective of the RX Project is to develop a medical information system capable of accurately deriving knowledge of the course and consequences of treatment of chronic diseases from a large collection of stored patient records.

Computerized clinical databases and automated medical records systems have been under development throughout the world for at least a decade. Among the earliest of these endeavors was the ARAMIS Project, (American Rheumatism Association Medical Information System) under development at Stanford by Dr. James Fries and his colleagues since 1969. A prototype ambulatory records system was generalized in the early 1970's by Prof. Gio Wiederhold and Stephen Weyl in the form of a Time-Oriented Database (TOD) System. The TOD System, run on the IBM 370/3033 at the Stanford Center for Information Processing (SCIP), now supports the ARAMIS Project as well as a host of other chronic disease databases which store patient data gathered at many institutions nation-wide. At the present time ARAMIS contains records of over 14,000 patients with a variety of rheumatologic diagnoses. Over 62,000 patient visits have been recorded, accounting for 50,000 patient-years of observation.

The fundamental objective of ARAMIS, the other TOD research groups, and all other clinical data bank researchers is to use the raw data which has been gathered by clinical observation in order to study the evolution and medical management of chronic diseases. Unfortunately, the process of reliably deriving knowledge from raw data has proven to be refractory to existing techniques because of problems stemming from the complexity of disease, therapy, and outcome definitions; the complexity of time relationships; complex causal relationships creating strong sources of bias; and problems of missing and outlying data.

A major objective of the RX Project is to explore the utility of symbolic computational methods and knowledge-based techniques at solving this problem of accurate knowledge inference from non-randomized, non-protocol patient records. A central component of RX is a knowledge base of medicine and statistics, organized as a hierarchy or taxonomic tree consisting of nodes with attached data and procedures. Nodes representing diseases and therapeutic regimens contain procedures which use a variety of time-dependent predicates to label patient records in the database, facilitating the retrieval of time-intervals of interest in the records. The database is then inverted so that each node or object in the knowledge base contains pointers to all time-intervals during which its definition is satisfied.

Nodes in the knowledge base also contain lists of other nodes which are causally related. These functional dependencies are used to infer causal pathways among nodes for purposes of selecting confounding variables which need to be controlled for in the study of a specific hypothesis. Causal pathways may also be used in an exploratory mode to assist in the discovery of new hypotheses.

To study a particular causal hypothesis the knowledge base also contains information on the applicability of various statistical procedures and procedures for applying them.

B. Medical Relevance and Collaboration

As a test bed for system development our focus of attention has been on the records of patients with systemic lupus erythematosus (SLE) contained in the Stanford portion of the ARAMIS Data Bank. SLE is a chronic rheumatologic disease with a broad spectrum of manifestations which can lead to death in the third decade of life. With many perplexing diagnostic and therapeutic dilemmas, it is a disease of considerable medical interest.

In the future we anticipate possible collaborations with other project users of the TOD System such as the National Stroke Data Bank, the Northern California Oncology Group, and the Stanford Divisions of Oncology and of Radiation Therapy.

The RX Project is a new research effort only in existence for about two years, and, hence the project is still in a developmental stage. The primary issues being addressed at this stage are those concerned with the specifics of knowledge representation.

We believe that this research project is broadly applicable to the entire gamut of chronic diseases which constitute the bulk of morbidity and mortality in the United States. Consider five major diagnostic categories which are responsible for approximately two thirds of the two million deaths per year in the United States: myocardial infarction, stroke, cancer, hypertension, and diabetes. Therapy for each of these diagnoses is fraught with controversy concerning the balance of benefits versus costs.

- 1) Myocardial Infarction: Indications for and efficacy of coronary artery bypass graft vs. medical management alone. Indications for long-term antiarrhythmics ... long-term anticoagulants. Benefits of cholesterol-lowering diets, exercise, etc.
- 2) Stroke: Efficacy of long-term anti-platelet agents, long-term anticoagulation. Indications for revascularization.
- 3) Cancer: Relative efficacy of radiation therapy, chemotherapy, surgical excision - singly or in combination. Optimal frequency of screening procedures. Prophylactic therapy.
- 4) Hypertension: Indications for therapy. Efficacy versus adverse effects of chronic antihypertensive drugs. Role of various diagnostic tests such as renal arteriography in work-up.
- 5) Diabetes: Influence of insulin administration on microvascular complications. Role of oral hypoglycemics.

Despite the expenditure of billions of dollars over recent years for randomized controlled trials (RCT's) designed to answer these and other questions, answers have been slow in coming. RCT's are expensive of funds and personnel. The therapeutic questions in clinical medicine are too numerous for each to be addressed by its own series of RCT's.

On the other hand, the data regularly gathered in patient records in the course of the normal performance of health care delivery is a rich and largely underutilized resource. The ease of accessibility and manipulation of these data afforded by computerized clinical data banks holds out the possibility of a major new resource for acquiring knowledge on the evolution and therapy of chronic diseases.

The goal of the research which we are pursuing on SUMEX is to increase the reliability of knowledge derived from clinical data banks with the hope of providing a new tool for augmenting knowledge of diseases and therapies as a supplement to knowledge derived from formal prospective clinical trials. Furthermore, the incorporation of knowledge from both clinical data banks and other sources into a uniform knowledge base should increase the ease of access by individual clinicians to this knowledge and thereby facilitate both the practice of medicine as well as the investigation of human disease processes.

C. Highlights of Research Progress

1. 1 July 1980 to 1 May 1981

Our predominant objective was to detail the overall conceptual framework for the knowledge base and to develop the extensive computational machinery necessary for retrieving, analyzing, and displaying defined time-intervals within patient records.

The RX Knowledge Base (KB):

The central component of RX is a knowledge base of medicine and statistics, organized as a frame-based, taxonomic tree consisting of units with attached data and procedures. Units representing diseases and therapies contain procedures which use a variety of time-dependent predicates to label the patient records, facilitating the retrieval of time-intervals of interest in the records. Other units representing statistical techniques are used to map hypotheses onto study designs and event definitions. Implementing the algorithms and data structures of this KB was one of the major tasks of the current year.

At the current time the RX KB contains about 200 units of which 75 contain definitions and other relevant information pertaining to disease courses, effects of drugs, lab values, etc. This information comprises a small subset of medical knowledge dealing with some of the signs and symptoms of systemic lupus erythematosus (SLE) as well as the effects and indications of some drugs used for this disease. Other units contain machine-readable knowledge of statistical techniques needed for testing entered hypotheses. There are approximately 40 time-dependent functions used to map from the database values onto defined units.

The entire RX system currently contains approximately 400 INTERLISP functions accounting for 150 disk pages of code. The KB is about 60 disk pages. One disk page = 512 words * 36 bits per word. Also one disk page = approx. 1.5 typed pages on 8.5 by 11.5 inch paper.

Statistical Interfaces:

Once the relevant episodes have been defined and retrieved from the database they must be analyzed statistically. To do this we have recently adopted the IDL or Interactive Data-Analysis Language package developed at the Xerox Palo Alto Research Corp. IDL is a matrix manipulation language similar to APL and is built upon INTERLISP as is RX itself. The use of IDL for statistical analysis confers a tremendous advantage in that analyses are now highly interactive. IDL has completely supplanted our use of SPSS.

Time-Oriented Graphics Package:

This package enables data on an individual patient to be graphed over time, either linearly by visit or by calendar time with a "telescoping" capability. The program overlays graphs of both point data and data represented as episodes.

Study Editor:

Dr. Jerrold Kaplan, a research associate affiliated with the project, has implemented an additional package of programs which display to the clinician user those decisions which have been made by the knowledge base concerning which statistical techniques are to be employed, which variables are to be controlled for, and which time intervals are to be excluded. This affords the user with a means for seeing a sketch of the study plan before it is executed, and enables him to modify that plan.